

Improved Bounds on RIP for Generalized Orthogonal Matching Pursuit

Siddhartha Satpathi¹, Rajib Lochan Das² and Mrityunjay Chakraborty³

Department of Electronics and Electrical Communication Engineering

Indian Institute of Technology, Kharagpur, INDIA

E.Mail :¹sidd.piku@gmail.com, ²rajob.das.iit@gmail.com, ³mrityun@ece.iitkgp.ernet.in

Abstract—Generalized Orthogonal Matching Pursuit (gOMP) is a natural extension of OMP algorithm where unlike OMP, it may select $N(\geq 1)$ atoms in each iteration. In this paper, we demonstrate that gOMP can successfully reconstruct a K -sparse signal from a compressed measurement $\mathbf{y} = \Phi\mathbf{x}$ by K^{th} iteration if the sensing matrix Φ satisfies restricted isometry property (RIP) of order NK where $\delta_{NK} < \frac{\sqrt{N}}{\sqrt{K+2\sqrt{N}}}$. Our bound offers an improvement over the very recent result shown in [1]. Moreover, we present another bound for gOMP of order $NK+1$ with $\delta_{NK+1} < \frac{\sqrt{N}}{\sqrt{K+\sqrt{N}}}$ which exactly relates to the near optimal bound of $\delta_{K+1} < \frac{1}{\sqrt{K+1}}$ for OMP ($N=1$) as shown in [2].

I. INTRODUCTION

Compressed sensing or compressive sampling (CS) [3]–[5] is a powerful technique to represent signals at a sub-Nyquist sampling rate while retaining the capacity of perfect (or near perfect) reconstruction of the signal, provided the signal is known to be sparse in some domain. In last few years, the CS technique has attracted considerable attention from across a wide array of fields like applied mathematics, statistics, and engineering, including signal processing areas such as MR imaging, speech processing, analog to digital conversion etc. Let a real valued, band-limited signal be sampled following Nyquist sampling rate over a finite observation interval, generating a $n \times 1$ signal vector $\mathbf{u} = (u_1, u_2, \dots, u_n)^T$. The vector \mathbf{u} is known to be K -sparse under some transform domain

$$\mathbf{u} = \Psi\mathbf{x}$$

where Ψ is $n \times n$ transform matrix and \mathbf{x} is the corresponding n dimensional transform coefficient vector that is approximated with at most K non-zero entries. Suppose that the signal \mathbf{u} is converted to a lower dimension (m) via linear random projection

$$\mathbf{y} = \mathbf{A}\mathbf{u}$$

where \mathbf{y} is observation vector with $m \ll n$ and \mathbf{A} is a flat $m \times n$ random matrix. According to the CS theory, it is then possible to reconstruct the signal \mathbf{u} exactly from a very limited number of measurements $M = \mathcal{O}(K \log_e n)$. Therefore, CS framework results in a potential challenge in reconstructing a K -sparse signal from an under determined system equation

$$\mathbf{y} = \Phi\mathbf{x}$$

where $\Phi = \mathbf{A}\Psi$ is a $m \times n$ dimensional sensing matrix.

Under the K -sparse assumption \mathbf{x} can be reconstructed by solving the following l_0 minimization problem

$$\min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{x}\|_0 \text{ subject to } \mathbf{y} = \Phi\mathbf{x}. \quad (1)$$

[Note that uniqueness of the K -sparse solution requires every $2K$ column of Φ to be linearly independent.] The above l_0 minimization problem provides the sparsest solution for \mathbf{x} . However, the l_0 minimization problem is a non-convex problem and is NP-hard. The feasible practical algorithm for this inverse problem may be broadly classified into two categories, namely convex relaxation and greedy pursuits.

1) *Convex Relaxation*: This approach translates the non-convex l_0 problem into relaxed convex problem using its closest convex l_1 norm. This imposes ‘‘Restricted Isometry Property (RIP)’’ condition of appropriate order on Φ as defined below.

Definition 1. A matrix $\Phi^{m \times n}$ satisfies RIP of order K if there exists a constant $\delta \in (0, 1)$ for all index set $I \subset \{1, 2, \dots, N\}$ with $|I| \leq K$ such that

$$(1 - \delta)\|\mathbf{q}\|_2^2 \leq \|\Phi_I \mathbf{q}\|_2^2 \leq (1 + \delta)\|\mathbf{q}\|_2^2. \quad (2)$$

The RIP constant δ_K is defined as the smallest value of all δ for which the RIP is satisfied.

There are three main directions under this category, namely the basis pursuit (BP) [6], the basis pursuit de-

noising (BPDN) [7] and the LASSO [8]. The reconstruction problem is formulated under them as,

1. *BP*: $\min_{\mathbf{x} \in \mathbb{R}^N} \|\mathbf{x}\|_1 \text{ subject to } \mathbf{y} = \Phi \mathbf{x}$
2. *BPDN*: $\min_{\mathbf{x} \in \mathbb{R}^N} \lambda \|\mathbf{x}\|_1 + \|\mathbf{r}\|_2^2 \text{ s.t. } \mathbf{r} = \mathbf{y} - \Phi \mathbf{x}$
3. *LASSO*: $\min_{\mathbf{x} \in \mathbb{R}^N} \|\mathbf{y} - \Phi \mathbf{x}\|_2^2 \text{ s.t. } \|\mathbf{x}\|_1 \leq \epsilon$

The BP problem can be solved by standard polynomial time algorithms of linear programming. The exact K -sparse signal reconstruction by BP algorithm based on RIP was first investigated in [9] with the following bound on δ : $\delta_K + \delta_{2K} + \delta_{3K} < 1$. Later the bound was refined as $\delta_{2K} < \sqrt{2} - 1$ [10], $\delta_{1.75K} < \sqrt{2} - 1$ [11] and $\delta_{2K} < 0.4652$ [12]. The BPDN and LASSO problem can be solved by efficient quadratic programming (QP) like primal-dual interior method. However, the regularization parameters λ and ϵ play a crucial role in the performance of these algorithms. The convex relaxation technique provides uniform guarantee for sparse recovery. However, the complexity of ℓ_1 minimization technique is large enough ($\mathcal{O}(n^3)$) for some applications (e.g. real time video processing).

2) *Greedy Pursuits*: This approach recovers the K -sparse signal by iteratively constructing the support set of the sparse signal (*i.e.* index of non-zero elements in the sparse vector). At each iteration, it updates its support set by appending the index of one or more columns (called atoms) of the matrix Φ (often called dictionary) by some greedy principles. This category includes algorithms like orthogonal matching pursuit (OMP) [13], generalized orthogonal matching pursuit (gOMP) [1], [14], orthogonal least square (OLS) [15], compressive sampling matching pursuit (CoSaMP) [16], subspace pursuit (SP) [17] and so on. These algorithms offer very fast convergence rate with high accuracy in reconstruction performance, but they lack proper theoretical convergence guaranty. Among these greedy algorithms, OMP is widely used because of its simplicity. The theoretical guaranty of OMP algorithm for an exact recovery of the sparse signal under a $K + 1^{th}$ order RIP condition on Φ is improved in the following way: $\delta_{K+1} < \frac{1}{3\sqrt{K}}$ in [18], $\delta_{K+1} < \frac{1}{1+2\sqrt{K}}$ in [19], $\delta_{K+1} < \frac{1}{\sqrt{2K}}$ in [14] and $\delta_{K+1} < \frac{1}{\sqrt{K+1}}$ in [2], [20].

A. Our contribution in this paper

In this paper, we have analyzed the theoretical performance of gOMP algorithm in a different approach and our theoretical result improves the bound on RIP of order NK from $\delta_{NK} < \frac{\sqrt{N}}{\sqrt{K+3\sqrt{N}}}$ [1] to $\delta_{NK} < \frac{\sqrt{N}}{\sqrt{K+2\sqrt{N}}}$. we have also presented another approach which results in

a RIP bound of order $NK + 1$ with $\delta_{NK+1} < \frac{\sqrt{N}}{\sqrt{K+\sqrt{N}}}$. Finally, we have discussed the theoretical performance of this algorithm under noisy measurement and proposed a bound on signal to noise ratio ($\text{SNR} = \frac{\|\mathbf{y}\|_2}{\|\mathbf{n}\|_2}$) for correct reconstruction of support set.

B. Organization of the paper

Rest of the paper is organized as follows. Next section presents the notations used in this paper and a brief review of OMP and gOMP algorithms. In section III, theoretical analysis of gOMP algorithm for noiseless observations is presented. In section IV, analysis of this algorithm in presence of noise is provided. Discussion is presented in section V and conclusions are drawn in section VI.

II. NOTATIONS AND A BRIEF REVIEW OF OMP AND gOMP ALGORITHMS

A. Notations

The following notations will be used in this paper. Let the columns of Φ matrix be called as atoms where $\Phi = [\phi_1 \phi_2 \phi_3 \dots \phi_n]$. The matrix Φ_A represents the sub-matrix of Φ with columns indexed by the elements present in set A . Similarly \mathbf{x}_A represents the sub-vector of \mathbf{x} with elements whose indices are given in set A . T is the true support set of \mathbf{x} and Λ^k is the estimated support set after k iterations of algorithm. $\Phi_{\Lambda^k}^\dagger = (\Phi_{\Lambda^k}^T \Phi_{\Lambda^k})^{-1} \Phi_{\Lambda^k}^T$ is the pseudo-inverse of Φ_{Λ^k} . Here we assume that Φ_{Λ^k} has full column rank ($\Lambda^k < m$). $\mathbf{P}_{\Lambda^k} = \Phi_{\Lambda^k} \Phi_{\Lambda^k}^\dagger$ is the projection operator onto column space of Φ_{Λ^k} and $\mathbf{P}_{\Lambda^k}^\perp = \mathbf{I} - \mathbf{P}_{\Lambda^k}$ is the projection operator upon the rejection space of $\text{span}(\Lambda^k)$. $\mathbf{A}_{\Lambda^k} = \mathbf{P}_{\Lambda^k}^\perp \Phi$ is a matrix obtained by orthogonalizing (projecting onto rejection space) the columns of Φ against $\text{span}(\Phi_{\Lambda^k})$.

For referring to previous results we use the following notation. Suppose an equation follows from the result of Lemma 1 then L1 is mentioned at the top of the inequality/equality like $\stackrel{L1}{>}$. Similarly if an equation follows from another equation or definition or theorem then it is mentioned as $\stackrel{(1)}{=}$ or $\stackrel{D1}{=}$ or $\stackrel{T1}{>}$ respectively.

B. A brief review of OMP and gOMP algorithms

The algorithm is presented in Table 1. The OMP algorithm starts with an empty support set Λ^0 and keep selecting a single atom in every iteration based on highest correlation with residual signal \mathbf{r}^{k-1} until the support set is full with the index of K atoms. At k^{th} iteration, the residual signal \mathbf{r}^k is updated using the difference between signal \mathbf{y} and its orthogonal projection

TABLE I
OMP ALGORITHM

Input: measurement $\mathbf{y} \in \mathbb{R}^m$, sensing matrix $\Phi^{m \times n}$
Initialization: counter $k=0$, residue $\mathbf{r}^0=\mathbf{y}$, estimated support set $\Lambda^k = \emptyset$
While $k < K$ and $\ \mathbf{r}^k\ _2 < \ \mathbf{r}^{k-1}\ _2$ $k=k+1$ Identification: $h^k = \arg \max_j \langle \mathbf{r}^{k-1}, \Phi_j \rangle $ Augment: $\Lambda^k = \Lambda^{k-1} \cup \{h^k\}$ Estimate: $\mathbf{x}_{\Lambda^k} = \arg \min_{z: \text{supp}(z) \in \Lambda^k} \ \mathbf{y} - \Phi_{\Lambda^k} z\ _2$ Update: $\mathbf{r}^k = \mathbf{y} - \Phi_{\Lambda^k} \mathbf{x}_{\Lambda^k}$ End While Output: $\mathbf{x} = \arg \min_{z: \text{supp}(z) \in \Lambda^K} \ \mathbf{y} - \Phi_{\Lambda^K} z\ _2$
In gOMP algorithm the Identification step is only different. We select a vector $\mathbf{h}^k \in \mathbb{R}^N$ which has N largest entries in $ \Phi^T \mathbf{r}^{k-1} $. ($NK < m$)

on the subspace spanned by the atoms corresponding to the current support set Λ^k . Generalized OMP algorithm is very similar to OMP where N largest correlated atoms are selected in each step. This simple modification in identification step results in improved reconstruction performance for K -sparse signal [1].

III. ANALYSIS OF gOMP

To analyse gOMP algorithm we use some commonly used properties of RIP as summarized in *Lemma 1*.

Lemma 1. (*Lemma 1 in [17] [16]*)

- a) $\delta_{K_1} < \delta_{K_2} \forall K_1 < K_2$ (*monotonicity*)
- b) $(1 - \delta_{|I|})\|\mathbf{q}\|_2 \leq \|\Phi_I^T \Phi_I \mathbf{q}\|_2 \leq (1 + \delta_{|I|})\|\mathbf{q}\|_2$
- c) $\|\Phi_I^T \mathbf{q}\|_2 < \sqrt{1 + \delta_{|I|}}\|\mathbf{q}\|_2$
- d) $\langle \Phi_I \mathbf{q}, \Phi_J \mathbf{p} \rangle \leq \|\mathbf{q}\|_2 \|\Phi_I^T \Phi_J \mathbf{p}\|_2 <$

$\delta_{|I|+|J|}\|\mathbf{p}\|_2\|\mathbf{q}\|_2$
for $I, J \subset \{1, 2, \dots, n\}$, $|I|, \mathbf{q} \in \mathbb{R}^I$ and $\mathbf{p} \in \mathbb{R}^J$

Note that, the algorithm can reconstruct a K -sparse signal by K^{th} iterations if atleast one correct index is chosen in each iteration. Now, let in $k+1^{th}$ iteration $\beta_i^k = \langle \Phi_i, \mathbf{r}^k \rangle$ for $i \in T$ and $\alpha_j^k = \langle \Phi_j, \mathbf{r}^k \rangle$ for $j \notin T$ where β_i^k 's and α_j^k 's are arranged in descending order. So $\beta_1^k > \beta_2^k > \dots > \beta_N^k$ are N largest correlations in support set and similarly $\alpha_1^k > \alpha_2^k > \dots > \alpha_N^k$ are N largest correlations of incorrect indices. Now if we ensure that $\beta_1^k > \alpha_N^k$ then atleast β_1^k will appear in the overall N largest correlated atoms which are selected. Hence, we find the lower bound of β_1^k and upper bound of α_N^k and compare them. In this paper, we propose two RIP bounds which are presented as Theorem 1 and Theorem 2.

Theorem 1. *gOMP can recover \mathbf{x} exactly when Φ*

satisfies RIP of order NK with

$$\delta_{NK} < \frac{\sqrt{N}}{\sqrt{K} + 2\sqrt{N}}$$

Proof. To start with we use the same upper bound on α_N^k as presented in [1]. Interested readers may refer [1] for proof of the following lemma.

Lemma 2. (*Lemma 3.6 in [1]*)

$$\alpha_N^k < \frac{\delta_{NK}}{1 - \delta_{NK}} \frac{\|\mathbf{x}_{T-\Lambda^k}\|_2}{\sqrt{N}}$$

Now we go about finding a better bound on β_1^k . We observe that $\mathbf{r}^k = \mathbf{P}_{\Lambda^k}^\perp \mathbf{y}$ and $\beta_1^k = \|\Phi_T^T \mathbf{r}^k\|_\infty$ and for any $i \in \Lambda^k$

$$\langle \phi_i, \mathbf{r}^k \rangle = \langle \phi_i, \mathbf{P}_{\Lambda^k}^\perp \mathbf{r}^k \rangle = \langle \mathbf{P}_{\Lambda^k}^\perp \phi_i, \mathbf{r}^k \rangle = 0 \quad (3)$$

So,

$$\begin{aligned} \|\Phi_T^T \mathbf{r}^k\|_\infty &> \frac{1}{\sqrt{K}} \|\Phi_T^T \mathbf{r}^k\|_2 \quad (\text{as } |T| = K) \\ &\stackrel{(3)}{=} \frac{1}{\sqrt{K}} \|\Phi_{T-\Lambda^k}^T \mathbf{r}^k\|_2 = \frac{1}{\sqrt{K}} \|\Phi_{T-\Lambda^k}^T \mathbf{P}_{\Lambda^k}^\perp \mathbf{y}\|_2 \\ &= \frac{1}{\sqrt{K}} \|\Phi_{T-\Lambda^k}^T (\mathbf{P}_{\Lambda^k}^\perp)^T \mathbf{P}_{\Lambda^k}^\perp \mathbf{y}\|_2 \quad (\text{as } \mathbf{P} = \mathbf{P}^T \& \mathbf{P} = \mathbf{P}^2) \\ &= \frac{1}{\sqrt{K}} \|(\mathbf{P}_{\Lambda^k}^\perp \Phi_{T-\Lambda^k})^T \mathbf{P}_{\Lambda^k}^\perp \Phi_T \mathbf{x}_T\|_2 \\ &\stackrel{(3)}{=} \frac{1}{\sqrt{K}} \|(\mathbf{P}_{\Lambda^k}^\perp \Phi_{T-\Lambda^k})^T \mathbf{P}_{\Lambda^k}^\perp \Phi_{T-\Lambda^k} \mathbf{x}_{T-\Lambda^k}\|_2 \end{aligned} \quad (4)$$

Now, to proceed further we require the following lemma.

Lemma 3. (*Extension of lemma 3.2 from [18]*)

$\mathbf{A}_{I_1}^{m \times n}$ satisfies modified RIP of $\frac{\delta_{|I_1|+|I_2|}}{1 - \delta_{|I_1|+|I_2|}}$

$$(1 - \frac{\delta_{|I_1|+|I_2|}}{1 - \delta_{|I_1|+|I_2|}}) \|\mathbf{u}\|_2^2 < \|\mathbf{A}_{I_1} \mathbf{u}\|_2^2 < (1 + \delta_{|I_1|+|I_2|}) \|\mathbf{u}\|_2^2$$

and also

$$(1 - (\frac{\delta_{|I_1|+|I_2|}}{1 - \delta_{|I_1|+|I_2|}})^2) \|\Phi \mathbf{u}\|_2^2 < \|\mathbf{A}_{I_1} \mathbf{u}\|_2^2 < \|\Phi \mathbf{u}\|_2^2$$

where $\mathbf{u} \in \mathbb{R}^n$, $I_1, I_2 \in \{1, \dots, n\}$ $\text{supp}(\mathbf{u}) \in I_2$ and $I_1 \cap I_2 = \emptyset$

Proof: In Appendix A

$$\begin{aligned} \text{Now Let } \mathbf{x}' &= \begin{bmatrix} \mathbf{x}_{T-\Lambda^k} \\ 0 \end{bmatrix}, \mathbf{x}' \in \mathbb{R}^n. \text{ So, } \mathbf{A}_{\Lambda^k} \mathbf{x}' = \\ \mathbf{P}_{\Lambda^k}^\perp \Phi \mathbf{x}' &= \mathbf{P}_{\Lambda^k}^\perp \begin{bmatrix} \Phi_{T-\Lambda^k} & \Phi_{(T-\Lambda^k)^c} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{T-\Lambda^k} \\ 0 \end{bmatrix} = \\ \mathbf{P}_{\Lambda^k}^\perp \Phi_{T-\Lambda^k} \mathbf{x}_{T-\Lambda^k} &\text{ and also } \mathbf{P}_{\Lambda^k}^\perp \Phi_T \mathbf{x}_T = \\ \mathbf{P}_{\Lambda^k}^\perp \begin{bmatrix} \Phi_{T-\Lambda^k} & \Phi_{(T \cap \Lambda^k)} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{T-\Lambda^k} \\ \mathbf{x}_{T \cap \Lambda^k} \end{bmatrix} &\stackrel{(3)}{=} \\ \mathbf{P}_{\Lambda^k}^\perp \Phi_{T-\Lambda^k} \mathbf{x}_{T-\Lambda^k}. &\text{ Hence we get} \end{aligned}$$

$$\begin{aligned} \mathbf{A}_{\Lambda^k} \mathbf{x}' &= \mathbf{P}_{\Lambda^k}^\perp \Phi \mathbf{x}' = \mathbf{P}_{\Lambda^k}^\perp \Phi_T \mathbf{x}_T \\ &= \mathbf{r}^k = \mathbf{P}_{\Lambda^k}^\perp \Phi_{T-\Lambda^k} \mathbf{x}_{T-\Lambda^k} \end{aligned} \quad (5)$$

Hence, from Lemma 3, with $I_1 = \Lambda^k$ and $I_2 = \text{supp}(\mathbf{x}') = T - \Lambda^k$ and $I_1 \cap I_2 = \emptyset, |I_1| + |I_2| = Nk + K - l$ where $T \cap \Lambda^k = l$ we get

$$\begin{aligned} \|\mathbf{A}_{\Lambda^k} \mathbf{x}'\|_2^2 &> (1 - \frac{\delta_{Nk+K-l}}{1 - \delta_{Nk+K-l}}) \|\mathbf{x}'\|_2^2 \\ &\stackrel{L1a}{>} (1 - \frac{\delta_{NK}}{1 - \delta_{NK}}) \|\mathbf{x}_{T-\Lambda^k}\|_2^2 \end{aligned} \quad (6)$$

Moreover,

$$\begin{aligned} \|\mathbf{A}_{\Lambda^k} \mathbf{x}'\|_2^2 &\stackrel{(5)}{=} \|\mathbf{P}_{\Lambda^k}^\perp \Phi_{T-\Lambda^k} \mathbf{x}_{T-\Lambda^k}\|_2^2 \\ &= \langle \mathbf{P}_{\Lambda^k}^\perp \Phi_{T-\Lambda^k} \mathbf{x}_{T-\Lambda^k}, \mathbf{P}_{\Lambda^k}^\perp \Phi_{T-\Lambda^k} \mathbf{x}_{T-\Lambda^k} \rangle \\ &= \langle (\mathbf{P}_{\Lambda^k}^\perp \Phi_{T-\Lambda^k})^T \mathbf{P}_{\Lambda^k}^\perp \Phi_{T-\Lambda^k} \mathbf{x}_{T-\Lambda^k}, \mathbf{x}_{T-\Lambda^k} \rangle \\ &< \|(\mathbf{P}_{\Lambda^k}^\perp \Phi_{T-\Lambda^k})^T \mathbf{P}_{\Lambda^k}^\perp \Phi_{T-\Lambda^k} \mathbf{x}_{T-\Lambda^k}\|_2 \|\mathbf{x}_{T-\Lambda^k}\|_2 \end{aligned} \quad (7)$$

Combining (6) and (7) we get

$$\begin{aligned} \|(\mathbf{P}_{\Lambda^k}^\perp \Phi_{T-\Lambda^k})^T \mathbf{P}_{\Lambda^k}^\perp \Phi_{T-\Lambda^k} \mathbf{x}_{T-\Lambda^k}\|_2 \\ > (1 - \frac{\delta_{NK}}{1 - \delta_{NK}}) \|\mathbf{x}_{T-\Lambda^k}\|_2 \end{aligned} \quad (8)$$

Therefore combining this result with (4) and (8) we get

$$\beta_1^k > \frac{1}{\sqrt{K}} (1 - \frac{\delta_{NK}}{1 - \delta_{NK}}) \|\mathbf{x}_{T-\Lambda^k}\|_2 \quad (9)$$

Making lower bound on β_1^k greater than upper bound of α_N^k (from Lemma 2) bring us to the result. \square

The next theorem states the other bound for gOMP success.

Theorem 2. *gOMP can recover \mathbf{x} exactly when Φ satisfies RIP of order $NK + 1$ with*

$$\delta_{NK+1} < \frac{\sqrt{N}}{\sqrt{K} + \sqrt{N}}$$

Proof. Let us begin by examining the residue \mathbf{r}^k in the $k+1^{th}$ iteration. In [2] it was shown that $\mathbf{r}^k \in \text{span}(\Phi_T)$ for OMP where estimated support set $\Lambda^k \subset T$. Now we show that in cases where Λ^k and T are in general modelled as shown in Fig.1, \mathbf{r}^k is indeed spanned by $\Phi_{T \cup \Lambda^k}$.

$$\begin{aligned} \mathbf{r}^k &= \mathbf{y} - \Phi_{\Lambda^k} \Phi_{\Lambda^k}^\dagger \mathbf{y} \\ &= \Phi_T \mathbf{x}_T - \mathbf{P}_{\Lambda^k} \mathbf{y} \\ &= \Phi_{T-\Lambda^k} \mathbf{x}_{T-\Lambda^k} + \Phi_{T \cap \Lambda^k} \mathbf{x}_{T \cap \Lambda^k} \\ &\quad - \mathbf{P}_{\Lambda^k} (\Phi_{T-\Lambda^k} \mathbf{x}_{T-\Lambda^k} + \Phi_{T \cap \Lambda^k} \mathbf{x}_{T \cap \Lambda^k}) \\ &= \Phi_{T-\Lambda^k} \mathbf{x}_{T-\Lambda^k} - \mathbf{P}_{\Lambda^k} \Phi_{T-\Lambda^k} \mathbf{x}_{T-\Lambda^k} \\ &= \Phi_{T-\Lambda^k} \mathbf{x}_{T-\Lambda^k} - \Phi_{\Lambda^k} \mathbf{z}_{\Lambda^k} \\ &= \Phi_{T \cup \Lambda^k} \mathbf{x}_{T \cup \Lambda^k}'' \end{aligned} \quad (10)$$

$$(11)$$

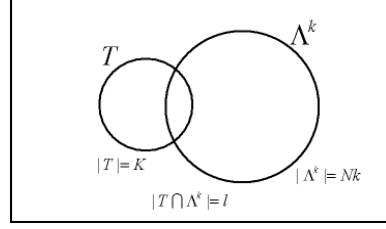


Fig. 1. Venn diagram for correct and estimated support set

where (10) follows from the fact that $\mathbf{P}_{\Lambda^k} \Phi_{T-\Lambda^k} \mathbf{x}_{T-\Lambda^k} \in \text{span}(\Phi_{\Lambda^k})$ and it can be viewed as $\Phi_{\Lambda^k} \mathbf{z}_{\Lambda^k}$ where $\mathbf{z}_{\Lambda^k} \in \mathbb{R}^{\Lambda^k}$ and $\mathbf{z}_{\Lambda^k} = \Phi_{\Lambda^k}^\dagger \Phi_{T-\Lambda^k} \mathbf{x}_{T-\Lambda^k}$. So $\mathbf{x}_{T \cup \Lambda^k}''$ is a vector in $\mathbb{R}^{T \cup \Lambda^k}$. Observe that

$$\mathbf{x}_{T \cup \Lambda^k}'' = \begin{bmatrix} \mathbf{x}_{T-\Lambda^k} \\ \mathbf{z}_{\Lambda^k} \end{bmatrix} \quad (12)$$

Let W be the set of remaining incorrect indices over which α_i^k 's are chosen ($W \in (T \cup \Lambda^k)^c$). So,

$$\begin{aligned} \alpha_N^k &= \min(\langle \Phi_i, \mathbf{r}^k \rangle) \quad (i \in W) \\ &< \frac{\sum \alpha_i}{N} < \sqrt{\frac{\sum \alpha_i^2}{N}} \quad (\text{as } |W| = N) \\ &= \frac{\|\Phi_W^T \mathbf{r}^k\|_2}{\sqrt{N}} \\ &\stackrel{(11)}{=} \frac{1}{\sqrt{N}} \|\Phi_W^T \Phi_{T \cup \Lambda^k} \mathbf{x}_{T \cup \Lambda^k}''\|_2 \\ &\stackrel{L1d}{=} \frac{1}{\sqrt{N}} \delta_{N+Nk+K-l} \|\mathbf{x}_{T \cup \Lambda^k}''\|_2 \\ &\stackrel{L1a}{<} \frac{1}{\sqrt{N}} \delta_{NK+1} \|\mathbf{x}_{T \cup \Lambda^k}''\|_2 \end{aligned} \quad (13)$$

where (13) comes from the fact that $l \geq k$ and $k \leq K-1$.

Now for finding lower bound of β_1^k in terms of $\|\mathbf{x}_{T \cup \Lambda^k}''\|_2$ we proceed in this way.

$$\begin{aligned} \beta_1^k &= \|\Phi_T^T \mathbf{r}^k\|_\infty \\ &> \frac{1}{\sqrt{K}} \|\Phi_T^T \mathbf{r}^k\|_2 \quad (\text{as } |T| = K) \\ &= \frac{1}{\sqrt{K}} \|[\Phi_T \quad \Phi_{\Lambda^k-T}]^T \mathbf{r}^k\|_2 \\ &= \frac{1}{\sqrt{K}} \|\Phi_{T \cup \Lambda^k}^T \Phi_{T \cup \Lambda^k} \mathbf{x}_{T \cup \Lambda^k}''\|_2 \\ &\stackrel{L1b}{>} \frac{1}{\sqrt{K}} (1 - \delta_{Nk+K-l}) \|\mathbf{x}_{T \cup \Lambda^k}''\|_2 \\ &\stackrel{L1a}{>} \frac{1}{\sqrt{K}} (1 - \delta_{NK}) \|\mathbf{x}_{T \cup \Lambda^k}''\|_2 \end{aligned} \quad (14)$$

where (14) comes as $\Phi_{\Lambda^k-T}^T \mathbf{r}^k = 0$ which follows from (3). Now from (13) and (15) ensuring $\beta_1^k > \alpha_N^k$ gives us the result. \square

IV. ANALYSIS IN PRESENCE OF NOISE

In case of noise we can model the measurement as $\mathbf{y}' = \mathbf{y} + \mathbf{n} = \Phi \mathbf{x} + \mathbf{n}$ where \mathbf{n} is the added noise. We can show the performance of this algorithm in presence of noise in two ways. One is by finding the upper bound of reconstruction error energy $\|\mathbf{x} - \mathbf{x}_{\Lambda^k}\|_2$ (as presented in [1]) and other is by providing a condition for exact reconstruction subject to upper bound on measurement $\text{SNR} = \frac{\|\mathbf{y}\|_2}{\|\mathbf{n}\|_2}$.

Theorem 3. *If $k = K$ forms the stopping criterion in gOMP with Φ satisfying $\delta_{NK+K} < 1$ and $\delta_{NK} < \frac{\sqrt{N}}{\sqrt{K+2\sqrt{N}}}$ then $\|\mathbf{x} - \mathbf{x}_{\Lambda^k}\|_2 < C_{K_1} \|\mathbf{n}\|_2$ where*

$$C_{K_1} = \frac{(1-\delta_{NK})(\sqrt{N}(1+\delta_K) + \sqrt{K(1+\delta_N)(1+\delta_K)})}{(\sqrt{N} - (\sqrt{K} + 2\sqrt{N})\delta_{NK})\sqrt{1-\delta_{NK+K}}} + \frac{2}{\sqrt{1-\delta_{NK+K}}}$$

Proof: In Appendix B

Theorem 4. *If $k = K$ forms the stopping criterion in gOMP with Φ satisfying $\delta_{NK+K} < 1$ and $\delta_{NK+1} < \frac{\sqrt{N}}{\sqrt{K+\sqrt{N}}}$ then $\|\mathbf{x} - \mathbf{x}_{\Lambda^k}\|_2 < C_{K_2} \|\mathbf{n}\|_2$ where*

$$C_{K_2} = \frac{(\sqrt{N}(1+\delta_K) + \sqrt{K(1+\delta_N)(1+\delta_K)})}{(\sqrt{N} - (\sqrt{K} + \sqrt{N})\delta_{NK+1})\sqrt{1-\delta_{NK+K}}} + \frac{2}{\sqrt{1-\delta_{NK+K}}}$$

Proof: In Appendix C

The above bounds provide a estimate on upper bound on reconstruction energy. But they do not guarantee estimation of correct support set. In some cases it may happen that the reconstruction error energy is bounded but the chosen support set is completely different. This may prove expensive because in most cases knowledge of correct support set is more important than knowledge of exact value at that position. So easily verifiable bounds guaranteeing reconstruction of correct support set are necessary.

In communication we often judge the performance by the SNR of the received signal. Before applying sparse reconstruction we do not have the information of energy of vector \mathbf{x} . But we do have knowledge of energy of clean measurement from transmitter's end. Hence by calculating the SNR at the receiver's end we can have an idea whether a particular algorithm can be implemented for reconstruction or not. This can be a good measure of performance analysis for reconstruction algorithms. So we present a bound on $\frac{\|\mathbf{y}\|_2}{\|\mathbf{n}\|_2}$ for which correct support set is estimated. Before stating the theorem let us analyse the assumption made: $|\mathbf{x}_i| > \frac{|\mathbf{x}_j|}{\gamma} \forall i, j \in T$. This implies that all non zero values of \mathbf{x} are bounded within some ratio of the maximum. We see that the sparse systems are modelled by setting the values of elements in \mathbf{x}

below some threshold as zero. Hence this assumption can always be made. If suppose \mathbf{x} has a non-zero value below $\frac{|x_{max}|}{\gamma}$ then it can be modelled as a $K-1$ sparse system by setting that value to zero without affecting the output much.

Theorem 5. *If measurement $\mathbf{y} = \Phi \mathbf{x}$ is corrupted with noise \mathbf{n} then gOMP algorithm can still recover the true support of \mathbf{x} provided $\frac{\|\mathbf{y}\|_2}{\|\mathbf{n}\|_2} > C_{K_3}$ where*

$$C_{K_3} = \frac{\sqrt{K(1+\delta_N)} + \sqrt{N(1+\delta_K)}}{\frac{\sqrt{N}(1-\delta_K)(1-2\delta_{NK})}{\gamma\sqrt{(1+\delta_K)(1-\delta_{NK})^2}} - \sqrt{K(1+\delta_N)}}$$

Proof. At first let us make use of the assumption and provide a result which would be used in subsequent proof

Lemma 4. *With usual notations we see that*

$$\|\Phi_{T-\Lambda^k} \mathbf{x}_{T-\Lambda^k}\|_2 > \sqrt{\frac{(K-l)(1-\delta_K)}{K(1+\delta_K)}} \frac{\|\Phi_T \mathbf{x}_T\|_2}{\gamma}$$

proof: In Appendix D

Let us again compute the bounds on α_N^k and β_1^k

$$\begin{aligned} \alpha_N^k &< \frac{\|\Phi_W^T \mathbf{P}_{\Lambda^k}^\perp \mathbf{y}'\|_2}{\sqrt{N}} \\ &\stackrel{L1c}{<} \frac{\sqrt{1+\delta_N} \|\mathbf{P}_{\Lambda^k}^\perp \mathbf{y}'\|_2}{\sqrt{N}} \\ &< \frac{\sqrt{1+\delta_N} \|\mathbf{y}'\|_2}{\sqrt{N}} \\ &< \frac{\sqrt{1+\delta_N} (\|\Phi_T \mathbf{x}_T\|_2 + \|\mathbf{n}\|_2)}{\sqrt{N}} \end{aligned} \quad (16)$$

$$\begin{aligned}
\beta_1^k &> \frac{1}{\sqrt{K-l}} \|\Phi_{T-\Lambda^k}^T \mathbf{P}_{\Lambda^k}^\perp (\Phi_{T-\Lambda^k} \mathbf{x}_{T-\Lambda^k} + \mathbf{n})\|_2 \\
&\stackrel{L1c}{>} \frac{1}{\sqrt{K-l}} \|\Phi_{T-\Lambda^k}^T \mathbf{P}_{\Lambda^k}^\perp \Phi_{T-\Lambda^k} \mathbf{x}_{T-\Lambda^k}\|_2 \\
&\quad - \frac{\sqrt{1+\delta_K}}{\sqrt{K}} \|\mathbf{n}\|_2 \\
&\stackrel{(7)}{>} \frac{\|\mathbf{A}_{\Lambda^k} \mathbf{x}'\|_2^2}{\sqrt{K-l} \|\mathbf{x}_{T-\Lambda^k}\|_2} - \frac{\sqrt{1+\delta_K}}{\sqrt{K}} \|\mathbf{n}\|_2 \\
&\stackrel{L3}{>} \frac{(1 - (\frac{\delta_{NK}}{1-\delta_{NK}})^2) \|\Phi_{T-\Lambda^k} \mathbf{x}_{T-\Lambda^k}\|_2^2}{\sqrt{K-l} \|\mathbf{x}_{T-\Lambda^k}\|_2} \\
&\quad - \frac{\sqrt{1+\delta_K}}{\sqrt{K}} \|\mathbf{n}\|_2 \\
&\stackrel{D1}{>} \frac{(1 - 2\delta_{NK}) \sqrt{1-\delta_K} \|\Phi_{T-\Lambda^k} \mathbf{x}_{T-\Lambda^k}\|_2}{\sqrt{K-l} (1-\delta_{NK})^2} \\
&\quad - \frac{\sqrt{1+\delta_K}}{\sqrt{K}} \|\mathbf{n}\|_2 \\
&\stackrel{L4}{>} \frac{(1 - 2\delta_{NK})(1 - \delta_K) \|\Phi_T \mathbf{x}_T\|_2}{(1 - \delta_{NK})^2 \gamma \sqrt{K}(1 + \delta_K)} - \frac{\sqrt{1+\delta_K}}{\sqrt{K}} \|\mathbf{n}\|_2
\end{aligned} \tag{17}$$

Making $\alpha_N^k < \beta_1^k$ ((16) < (17)) for correct choice of index we get the desired SNR bound. \square

V. DISCUSSION

The proposed bound in Theorem 1 is better than the one from [1] because while obtaining lower bound of β_1^k instead of applying successive inequalities we use a more direct inequality presented in Lemma 3 which leads us to a higher lower bound. This bound is also better than the bound proposed in [14] ($\delta_{NK} < \frac{\sqrt{K}}{(2+\sqrt{2})\sqrt{K}}$) for $N < 1.45K$. But according to [1] gOMP performs better than OMP for small values of N only.

It is difficult to compare the bounds presented in Theorem 1 and Theorem 2 since $\delta_{NK} < \delta_{NK+1}$ and $\frac{\sqrt{N}}{\sqrt{K+\sqrt{N}}} > \frac{\sqrt{N}}{\sqrt{K+2\sqrt{N}}}$. But intuitively we can see that bound on δ_{NK+1} is more optimal since it reduces to near optimal bound on OMP for special case of $N = 1$. The proposition on SNR seems to be a good approach since it is an easily measurable quantity and can be used in future research for comparing greedy algorithm's performance under noise.

VI. CONCLUSIONS

In this paper, we have given an elegant proof of the theoretical performance of gOMP algorithm. Our analysis improves the bound on RIP of order NK from $\delta_{NK} < \frac{\sqrt{N}}{\sqrt{K+3\sqrt{N}}}$ [1] to $\delta_{NK} < \frac{\sqrt{N}}{\sqrt{K+2\sqrt{N}}}$. In the same

paper, we have presented another bound of order $NK+1$ with RIP constant $\delta_{NK+1} < \frac{\sqrt{N}}{\sqrt{K+\sqrt{N}}}$. We have also presented improved theoretical performance of gOMP algorithm under noisy measurements.

APPENDIX A PROOF OF LEMMA 3

We know that

$$\begin{aligned}
\|\Phi \mathbf{u}\|_2^2 &= \|\mathbf{P}_{I_1} \Phi \mathbf{u}\|_2^2 + \|\mathbf{P}_{I_1}^\perp \Phi \mathbf{u}\|_2^2 \\
\Rightarrow \|\mathbf{A}_{I_1} \mathbf{u}\|_2^2 &= \|\Phi \mathbf{u}\|_2^2 - \|\mathbf{P}_{I_1} \Phi \mathbf{u}\|_2^2
\end{aligned} \tag{A.1}$$

Now $\langle \mathbf{P}_{I_1} \Phi \mathbf{u}, \Phi \mathbf{u} \rangle = (\Phi \mathbf{u})^T \mathbf{P}_{I_1}^T \Phi \mathbf{u} = (\Phi \mathbf{u})^T \mathbf{P}_{I_1}^T \mathbf{P}_{I_1} \Phi \mathbf{u} = \|\mathbf{P}_{I_1} \Phi \mathbf{u}\|_2^2$. Further we see that $\mathbf{P}_{I_1} \Phi \mathbf{u} \in \text{span}(\Phi_{I_1})$. So $\mathbf{P}_{I_1} \Phi \mathbf{u} = \Phi \mathbf{z}$ for some $\mathbf{z} \in \mathbb{R}^n$ with $\text{supp}(\mathbf{z}) \in I_1$.

$$\begin{aligned}
\frac{\|\mathbf{P}_{I_1} \Phi \mathbf{u}\|_2}{\|\Phi \mathbf{u}\|_2} &= \frac{\langle \mathbf{P}_{I_1} \Phi \mathbf{u}, \Phi \mathbf{u} \rangle}{\|\mathbf{P}_{I_1} \Phi \mathbf{u}\|_2 \|\Phi \mathbf{u}\|_2} = \frac{\langle \Phi \mathbf{z}, \Phi \mathbf{u} \rangle}{\|\Phi \mathbf{z}\|_2 \|\Phi \mathbf{u}\|_2} \\
&\stackrel{D1, L1d}{<} \frac{\delta_{|I_1|+|I_2|}}{\sqrt{1-\delta_{|I_1|}} \sqrt{1-\delta_{|I_2|}}} \\
&\stackrel{L1a}{<} \frac{\delta_{|I_1|+|I_2|}}{1-\delta_{|I_1|+|I_2|}}
\end{aligned} \tag{A.2}$$

So from (A.1) and (A.2) we get $\|\mathbf{A}_{I_1} \mathbf{u}\|_2^2 > (1 - (\frac{\delta_{|I_1|+|I_2|}}{1-\delta_{|I_1|+|I_2|}})^2) \|\Phi \mathbf{u}\|_2^2 > (1 - (\frac{\delta_{|I_1|+|I_2|}}{1-\delta_{|I_1|+|I_2|}})^2) (1 - \delta_{|I_2|}) \|\mathbf{u}\|_2^2 \stackrel{L1a}{>} (1 - \frac{\delta_{|I_1|+|I_2|}}{1-\delta_{|I_1|+|I_2|}}) \|\mathbf{u}\|_2^2$. Applying $\|\mathbf{P}_{I_1} \Phi \mathbf{u}\|_2^2 > 0$ in (A.1) the upper bound becomes $\|\mathbf{A}_{I_1} \mathbf{u}\|_2^2 < \|\Phi \mathbf{u}\|_2^2 < (1 + \delta_{|I_2|}) \|\mathbf{u}\|_2^2 \stackrel{L1a}{<} (1 + \delta_{|I_1|+|I_2|}) \|\mathbf{u}\|_2^2$.

APPENDIX B PROOF OF THEOREM 3

First we need to find the bounds on α_N^k and β_1^k in presence of noise in $k+1^{th}$ iteration

$$\begin{aligned}
\alpha_N^k &< \frac{\|\Phi_W^T \mathbf{r}^k\|_2}{\sqrt{N}} \\
&= \frac{1}{\sqrt{N}} \|\Phi_W^T (\mathbf{P}_{\Lambda^k}^\perp \mathbf{y} + \mathbf{P}_{\Lambda^k}^\perp \mathbf{n})\|_2 \\
&< \frac{1}{\sqrt{N}} (\|\Phi_W^T \mathbf{P}_{\Lambda^k}^\perp \mathbf{y}\|_2 + \|\Phi_W^T \mathbf{P}_{\Lambda^k}^\perp \mathbf{n}\|_2) \\
&\stackrel{L2, L1c}{<} \frac{1}{\sqrt{N}} \frac{\delta_{NK}}{1-\delta_{NK}} \|\mathbf{x}_{T-\Lambda^k}\|_2 + \frac{\sqrt{1+\delta_N}}{\sqrt{N}} \|\mathbf{P}_{\Lambda^k}^\perp \mathbf{n}\|_2 \\
&< \frac{1}{\sqrt{N}} \frac{\delta_{NK}}{1-\delta_{NK}} \|\mathbf{x}_{T-\Lambda^k}\|_2 + \frac{\sqrt{1+\delta_N}}{\sqrt{N}} \|\mathbf{n}\|_2
\end{aligned} \tag{B.1}$$

and

$$\begin{aligned}
\beta_1^k &> \frac{1}{\sqrt{K}} \|\Phi_{T-\Lambda^k}^T \mathbf{P}_{\Lambda^k}^\perp \mathbf{y}'\|_2 \\
&> \frac{1}{\sqrt{K}} \|\Phi_{T-\Lambda^k}^T \mathbf{P}_{\Lambda^k}^\perp \Phi_T \mathbf{x}_T\|_2 - \frac{1}{\sqrt{K}} \|\Phi_{T-\Lambda^k}^T \mathbf{P}_{\Lambda^k}^\perp \mathbf{n}\|_2 \\
&\stackrel{(9), L1c}{>} \frac{1}{\sqrt{K}} (1 - \frac{\delta_{NK}}{1 - \delta_{NK}}) \|\mathbf{x}_{T-\Lambda^k}\|_2 - \frac{\sqrt{1 + \delta_K}}{\sqrt{K}} \|\mathbf{P}_{\Lambda^k}^\perp \mathbf{n}\|_2 \\
&> \frac{1}{\sqrt{K}} (1 - \frac{\delta_{NK}}{1 - \delta_{NK}}) \|\mathbf{x}_{T-\Lambda^k}\|_2 - \frac{\sqrt{1 + \delta_K}}{\sqrt{K}} \|\mathbf{n}\|_2
\end{aligned} \tag{B.2}$$

Now at the end of algorithm it may happen that some incorrect atoms are chosen. Lets say this happens for the first time in the $p+1^{th}$ step. Then at this particular step (B.1)>(B.2). Which implies

$$\begin{aligned}
\|\mathbf{x}_{T-\Lambda^p}\|_2 &< \frac{(1 - \delta_{NK})(\sqrt{N(1 + \delta_K)} + \sqrt{K(1 + \delta_N)})}{(\sqrt{N} - (\sqrt{K} + 2\sqrt{N})\delta_{NK+1})} \\
&\quad \times \|\mathbf{n}\|_2
\end{aligned} \tag{B.3}$$

The error in reconstruction energy can be seen as

$$\begin{aligned}
\|\mathbf{x} - \mathbf{x}_{\Lambda^K}\|_2 &\stackrel{D1}{<} \frac{\|\Phi \mathbf{x} - \Phi \mathbf{x}_{\Lambda^K}\|_2}{\sqrt{1 - \delta_{NK+K}}} \\
&= \frac{\|\Phi \mathbf{x} - \Phi \Phi_{\Lambda^K}^\dagger \mathbf{y}'\|_2}{\sqrt{1 - \delta_{NK+K}}} \\
&= \frac{\|\mathbf{y}' - \mathbf{P}_{\Lambda^K} \mathbf{y}' - \mathbf{n}\|_2}{\sqrt{1 - \delta_{NK+K}}} \\
&< \frac{\|\mathbf{P}_{\Lambda^K}^\perp \mathbf{y}'\|_2 + \|\mathbf{n}\|_2}{\sqrt{1 - \delta_{NK+K}}} \\
&= \frac{\|\mathbf{r}^K\|_2 + \|\mathbf{n}\|_2}{\sqrt{1 - \delta_{NK+K}}} \\
&\leq \frac{\|\mathbf{r}^p\|_2 + \|\mathbf{n}\|_2}{\sqrt{1 - \delta_{NK+K}}} \text{ (as } \|\mathbf{r}^i\|_2 \leq \|\mathbf{r}^j\|_2 \text{ for } i > j) \\
&= \frac{\|\mathbf{P}_{\Lambda^p}^\perp (\Phi_T \mathbf{x}_T + \mathbf{n})\|_2 + \|\mathbf{n}\|_2}{\sqrt{1 - \delta_{NK+K}}} \\
&< \frac{\|\mathbf{P}_{\Lambda^p}^\perp \Phi_T \mathbf{x}_T\|_2 + \|\mathbf{P}_{\Lambda^p}^\perp \mathbf{n}\|_2 + \|\mathbf{n}\|_2}{\sqrt{1 - \delta_{NK+K}}} \\
&< \frac{\|\mathbf{P}_{\Lambda^p}^\perp \Phi_{T-\Lambda^p} \mathbf{x}_{T-\Lambda^p}\|_2 + 2\|\mathbf{n}\|_2}{\sqrt{1 - \delta_{NK+K}}} \\
&< \frac{\|\Phi_{T-\Lambda^p} \mathbf{x}_{T-\Lambda^p}\|_2 + 2\|\mathbf{n}\|_2}{\sqrt{1 - \delta_{NK+K}}} \\
&\stackrel{D1}{<} \frac{\sqrt{1 + \delta_K} \|\mathbf{x}_{T-\Lambda^p}\|_2 + 2\|\mathbf{n}\|_2}{\sqrt{1 - \delta_{NK+K}}}
\end{aligned} \tag{B.4}$$

By using (B.3) in (B.4) we get the desired bound.

APPENDIX C PROOF OF THEOREM 4

In this case we find the bounds on α_N^k and β_1^k in presence of noise similar to our proof on second bound of gOMP. Proceeding similar to (B.1) and (B.2) we get

$$\beta_1^k < \frac{1}{\sqrt{N}} \delta_{NK+1} \|\mathbf{x}_{T \cup \Lambda^k}''\|_2 + \frac{\sqrt{1 + \delta_N}}{\sqrt{N}} \|\mathbf{n}\|_2 \tag{C.1}$$

$$\beta_1^k > \frac{1 - \delta_{NK}}{\sqrt{K}} \|\mathbf{x}_{T \cup \Lambda^k}''\|_2 - \frac{\sqrt{1 + \delta_K}}{\sqrt{K}} \|\mathbf{n}\|_2 \tag{C.2}$$

So failure at $p+1^{th}$ step implies

$$\|\mathbf{x}_{T \cup \Lambda^p}''\|_2 < \frac{\sqrt{N(1 + \delta_K)} + \sqrt{K(1 + \delta_N)}}{(\sqrt{N} - (\sqrt{K} + \sqrt{N})\delta_{NK+1})} \|\mathbf{n}\|_2 \tag{C.3}$$

Now to get an upper bound in estimation error we proceed similarly as in (B.4)

$$\begin{aligned}
\|\mathbf{x} - \mathbf{x}_{\Lambda^K}\|_2 &< \frac{\sqrt{1 + \delta_K} \|\mathbf{x}_{T-\Lambda^p}\|_2 + 2\|\mathbf{n}\|_2}{\sqrt{1 - \delta_{NK+K}}} \\
&\stackrel{(12)}{<} \frac{\sqrt{1 + \delta_K} \|\mathbf{x}_{T \cup \Lambda^p}''\|_2 + 2\|\mathbf{n}\|_2}{\sqrt{1 - \delta_{NK+K}}} \\
&\stackrel{(C.3)}{<} C_{K_2} \|\mathbf{n}\|_2
\end{aligned} \tag{C.4}$$

APPENDIX D PROOF OF LEMMA 4

$$\begin{aligned}
\|\Phi_{T-\Lambda^k} \mathbf{x}_{T-\Lambda^k}\|_2 &\stackrel{D1}{>} \sqrt{1 - \delta_K} \|\mathbf{x}_{T-\Lambda^k}\|_2 \\
&> \sqrt{(1 - \delta_K)(K - l)} |\mathbf{x}_{min}| \\
&> \sqrt{\frac{(1 - \delta_K)(K - l)}{K}} \frac{\|\mathbf{x}_T\|_2}{\gamma} \\
&\stackrel{D1}{>} \sqrt{\frac{(1 - \delta_K)(K - l)}{K(1 + \delta_K)}} \frac{\|\Phi_T \mathbf{x}_T\|_2}{\gamma}
\end{aligned} \tag{D.1}$$

REFERENCES

- [1] J. Wang, S. Kwon, and B. Shim, "Generalized orthogonal matching pursuit," *IEEE Trans. on Signal Processing*, vol. 60, no. 12, pp. 6202–6216, dec. 2012.
- [2] J. Wang and B. Shim, "On the recovery limit of sparse signals using orthogonal matching pursuit," *IEEE Trans. on Signal Processing*, vol. 60, no. 9, pp. 4973–4976, sept. 2012.
- [3] D. Donoho, "Compressed sensing," *IEEE Trans. on Information Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [4] E. Candes, "Compressive sampling," in *Int. Congress of Mathematics*, vol. 3, 2006, pp. 1433–1452.
- [5] R. Baraniuk, "Compressed sensing," *IEEE Signal Processing Magazine*, vol. 25, pp. 21–30, 2007.
- [6] S. Chen and D. Donoho, "Basis pursuit," in *Proc. 28th Asilomar Conf. Signals, Syst. Comput.*, 1994, pp. 41–44.

- [7] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Scientif. Comput.*, vol. 20, no. 1, pp. 33–61, 1998.
- [8] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Royal. Statist. Soc. B.*, vol. 58, pp. 267–288, 1996.
- [9] J. R. E. Candès and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. on Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.
- [10] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," in *Compte Rendus de l'Academie des Sciences*, ser. I, no. 346, 2008, pp. 589–592.
- [11] G. X. J. Z. T. Tony Cai, "On recovery of sparse signals via l_1 minimization," *IEEE Trans. on Inf. Theory*, vol. 55, no. 7, pp. 3388–3397, 2009.
- [12] S. Foucart, "A note on guaranteed sparse recovery via l_1 -minimization," *Appl. Comput. Harmon. Anal.*, vol. 29, no. 1, pp. 97–103, 2010.
- [13] J. A. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. on Info. Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.
- [14] E. Liu and V. N. Temlyakov, "The orthogonal super greedy algorithm and applications in compressed sensing," *Inverse Problem*, vol. 58, no. 4, pp. 2040–2047, 2012.
- [15] S. A. B. S. Chen and W. . Luo, "Orthogonal least squares methods and their application to non-linear system identification," *Int. J. Contr.*, vol. 50, no. 5, pp. 1873–1896, 1989.
- [16] D. Needell and J. Tropp, "Cosamp : Iterative signal recovery from incomplete and inaccurate samples," *Appl. Comput. Harmon. Anal.*, vol. 26, pp. 301–321, 2009.
- [17] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Trans. on Inf. Theory*, vol. 55, no. 5, pp. 2230–2249, 2009.
- [18] M. A. Davenport and M. B. Wakin, "Analysis of orthogonal matching pursuit using the restricted isometry property," *IEEE Trans. on Information Theory*, vol. 56, no. 9, pp. 4395–4401, 2010.
- [19] S. Huang and J. Zhu, "Recovery of sparse signals using OMP and its variants: Convergence analysis based on RIP," *Inverse Problem*, vol. 27, no. 3, 2011.
- [20] Q. Mo and Y. Shen, "A remark on the restricted isometry property in orthogonal matching pursuit," *Information Theory, IEEE Transactions on*, vol. 58, no. 6, pp. 3654–3656, june 2012.